# LOCAL AND GLOBAL THEORY OF
# CONSTRAINED OPTIMIZATION

The results below, which deal with local and global optimization subject to equality and inequality constraints, will be discussed in class in the coming lectures, after the introduction of Fréchet differentials and derivatives (in short, F-differentials and F-derivatives).

First let me introduce the problem formulation for local optimization with equality constraints. The problem is formulated as one of minimization, but the development and all results below (in connection with this problem) apply verbatim to one of maximization.

Problem 1: Let $X$ and $Z$ be Banach spaces (B-spaces), and $f : X \to \mathbf{R}$, $H : X \to Z$ be two given functions. **Obtain** first-order necessary conditions for a vector $x^\circ \in X$ to minimize $f(x)$ locally subject to the equality constraint $H(x) = \theta$, where $\theta$ is the *zero* element of $Z$.

To solve Problem 1, we need the notion of a *regular point* of $H$:

Definition 1: Let $H$, as defined above, be continuously F-differentiable on an open domain $D \subset X$. Then, $x^\circ$ is a **regular point** of $H$, if the F-derivative $H'(x^\circ) : X \to Z$ is *onto*.

The following result (whose proof can be found in the text, p. 240) allows us to define a neighborhood of $x^\circ$ under the given equality constraint, so that local optimality can be tested.

**Theorem 1.** *(Generalized Inverse Function Theorem).* *Let $x^\circ \in D$ be as in Definition 1 above. Then, there exists a neighborhood $N(z^\circ)$ of $z^\circ = H(x^\circ)$, and a constant $K$, such that the equation $H(x) = z$ has a solution for every $z \in N(z^\circ)$,* **and** *the solution satisfies $\|x - x^\circ\| \leq K\|z - z^\circ\|$.*

The lemma below is the first step toward solving Problem 1.

**Lemma 1.** *In Problem 1, let $f$ and $H$ be continuously F-differentiable in an open neighborhood of $x^\circ \in X$, the locally minimizing solution, and further let $x^\circ$ be a regular point of $H$. Then,*

$$f'(x^\circ)h = 0 \quad \forall\, h \in X \quad \text{such that} \;\; H'(x^\circ)h = \theta$$

*Proof.* Let $T : X \to \mathbf{R} \times Z$ be defined by $T(x) := \big(f(x), H(x)\big)$. Assume that to the contrary $\exists\, h$ such that $H'(x^\circ)h = \theta$ but $f'(x^\circ)h \neq 0$. Then, $T'(x^\circ)$ is onto $\mathbf{R} \times Z$, since $H'(x^\circ)$ is onto $Z$ ($x^\circ$ being a regular point of the constraint). Hence, $x^\circ$ is a regular point of $T$, and by Theorem 1, the equation $T(x) = y$ has a solution when $y$ is in an open neighborhood of $y^\circ = \big(f(x^\circ), \theta\big) \in \mathbf{R} \times Z$ and $\|x - x^\circ\| \leq K\|y - y^\circ\|$. Choose $y = \big(f(x^\circ) - \delta, \theta\big)$ where $\delta > 0$ is sufficiently small. This then implies that $\exists\, x$ such that $T(x) = \big(f(x), H(x)\big) = \big(f(x^\circ) - \delta, \theta\big)$ and $\|x - x^\circ\| \leq K\delta$, which means that $x^\circ$ cannot be a local minimum.          ◇

The solution to Problem 1 (first-order necessary condition) is now given by the following theorem.

**Theorem 2.** *Consider Problem 1. Then, under the hypotheses of Lemma 1, $\exists\, z_o^* \in Z^*$ such that*

$$f'(x^\circ) + z_o^* H'(x^\circ) = \theta$$

*or equivalently, $x^°$ is a stationary point of the Lagrangian $L(x; z_o^*) := f(x) + z_o^* H(x)$.*

*Proof.* The result of Lemma 1 implies that $f'(x^°) \perp \mathcal{N}(H'(x^°))$, where $\mathcal{N}$ denotes the null space. But the range space of $H'(x^°)$, $\mathcal{R}(H(x^°))$, is closed (since it is onto), and hence $f'(x^°) \in \mathcal{R}(H(x^°)^*)$. This implies that $\exists z_o^* \in Z^*$ such that $f'(x^°) = -H'(x^°)^* z_o^* \equiv -z_o^* H'(x^°)$, where in the latter I have used the property of the adjoint $(< x, H'(x^°)^* z_o^* > = < H'(x^°)x, z_o^* >)$. This completes the proof. $\diamond$

**Remark 1.** *The results above are equally valid for locally maximizing solutions.*

The next problem and the ensuing result now involve **global minimization** for **convex functionals** under **convex constraints**.

Problem 2: Let $X$ be a vector space, and $Z$ be a normed linear space with positive cone $P$ with nonempty interior. Let $\Omega \subset X$ be a convex constraint set, and $f : X \to \mathbf{R}$ and $G : X \to Z$ be convex on $\Omega$. Let $\exists x_1 \in \Omega$ such that $G(x_1) < \theta$ (known as regularity of the inequality constraint). Finally, let the infimum of $f(x)$ over $x \in \Omega$ and under the inequality constraint $G(x) \leq \theta$ be finite, and denote it by $\mu_o$, that is

$$\mu_o := \inf_{x \in \Omega, G(x) \leq \theta} f(x) \tag{$*$}$$

**Obtain** the conditions for a vector $x^° \in X$ to minimize $f(x)$ globally over $\Omega$ subject to the inequality constraint $G(x) \leq \theta$.

The following theorem, known as **Karush-Kuhn-Tucker (KKT) Theorem**, provides the solution to Problem 2.

**Theorem 3.** *For Problem 2,*

(i) *There exists $z_o^* \in Z^*, z_o^* \geq \theta$, such that*

$$\mu_o = \inf_{x \in \Omega} L(x; z_o^*) \tag{$**$}$$

*where* $L(x; z^*) := f(x) + < G(x), z^* >$.

(ii) *If $(*)$ has a minimum, so does $(**)$, and the minimizing solutions are the same, i.e., $x^°$. Furthermore, $< G(x^°), z_o^* > = 0$.*

(iii) *If $x^°$ solves $(*)$, $L(x; z^*)$ has a saddle point over $(\Omega, Z_{\geq \theta}^*)$ at $(x^°, z_o^*)$, that is,*

$$L(x^°; z^*) \leq L(x^°; z_o^*) \leq L(x; z_o^*) \qquad \forall x \in \Omega, z^* \in Z_{\geq \theta}^* \tag{$***$}$$

$$\iff \quad \mu_o = \inf_{x \in \Omega} \sup_{z^* \in Z_{\geq \theta}^*} L(x; z^*) = \sup_{z^* \in Z_{\geq \theta}^*} \inf_{x \in \Omega} L(x; z^*) \equiv \sup_{z^* \in Z_{\geq \theta}^*} \psi(z^*) = L(x^°; z_o^*)$$

*where $\psi : Z^* \to \mathbf{R}$ is known as the dual function.*

Note that $(***)$ defines a zero-sum game, and clearly the following inequality holds: $\mu_o \geq \psi(z^*), \forall z^* \geq \theta$. This interpretation is valid even if Problem 2 does not have a solution, and even if convexity conditions do not hold. In such cases, the inequality could be a strict one, in which case we say that there is a **duality gap**. If $\exists z_o^* \geq \theta$ such that $\mu_o = \psi(z_o^*)$, then there is no duality gap.

There is also a **strong converse** to Theorem 3(iii), which directly links the existence of a SP solution to L to existence of a solution to Problem 2, without requiring convexity. Consider Problem 2, with convexity on $f, G, \Omega$ relaxed, but with $P$ being closed. Then, if $\exists (x^°, z_o^*)$ satisfying $(***)$, $x^°$ solves Problem 2.

$$\diamond \qquad \diamond \qquad \diamond$$